4

AD-A199 772

DIFFERENT ALGORITHMS FOR OBTAINING UPPER BOUNDS TO MULTIVARIATE NORMAL AREAS OUTSIDE OF ORIGIN CENTERED RECTANGLES USING JOINT MARGINAL PROBABILITIES

BY

DONALD R. HOOVER

TECHNICAL REPORT NO. 409

SEPTEMBER 27, 1988

Prepared Under Contract

N00014-86-K-0156 (NR-042-267)

For the Office of Naval Research

Herbert Solomon, Project Director

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

DIFFERENT ALGORITHMS FOR OBTAINING UPPER BOUNDS TO MULTIVARIATE NORMAL AREAS OUTSIDE OF ORIGIN CENTERED RECTANGLES USING JOINT MARGINAL PROBABILITIES

By

Donald R. Hoover

I. INTRODUCTION.

Upper bounds for multivariate normal probability areas outside of rectangles centered at the origin are of interest due to their applications in producing conservative unbiased simultaneous confidence intervals and hypothesis tests. Unfortunately, it is often not computationally practical to integrate multivariate normal area over dimensions $n > 4$ and obtain these probabilities exactly. The method commonly used to determine upper bounds for normal probabilities of these regions is based on the conservative assumption of independence as given in the following theorem suggested by Dunn (1958) and proved by Sidak (1967).

<u>Theorem I-1</u>. Let $\underset{\sim}{x} \sim N(0,\Sigma)$ and let A_i be the event $\{x_i \in (-c_i,c_i)\}$. Then

$$\Pr\{\bigcup_{i=1}^{n} A_i^c\} \le \sum_{i=1}^{n} \Pr\{A_i^c\} \Rightarrow \Pr\{\bigcap_{i=1}^{n} A_i\} \ge 1 - \sum_{i=1}^{n} \Pr\{A_i^c\} .$$

Since 1967, three different approaches have been taken which enable one to obtain lower upper bounds for the $\Pr\{\bigcap_{i=1}^{n} A_i\}$ than those given by Theorem I-1. These approaches require that Σ be known and integration

over joint m-variate densities be possible for some $m: 1 < m < n$ to obtain $Pr\{A_{i_1} \cup A_{i_2} \cup \cdots \cup A_{i_m}\}$. These three approaches are described in chapter II, and compared in chapter III. They are combined together in chapter IV to produce lower upper bounds for $Pr\{\bigcup_{i=1}^{n} A_i\}$ than are obtained from using any one approach alone. In chapter V examples of combining these methods are given and in chapter VI selected tables of upper bounds for $Pr\{\bigcup_{i=1}^{n} A_i\}$ given different combinations of the three approaches are shown. Chapter VII has a descriptive summary of the terminology used in this report and can be used for quick reference.

II. NEW APPROACHES FOR UPPER BOUNDS.

Approach I. Intersection Subtraction

This theorem was first proven for $m = 2$ by Hunter (1976) and later extended to $m > 2$ by Hoover (1987). The theorem is

Theorem A-I. Let $\underset{\sim}{x}$ be a $n \times 1$ vector distributed $N(0,\Sigma)$, $\underset{\sim}{y}$ be any permutation $\pi(\cdot)$ of the elements of $\underset{\sim}{x}$, and A_i be the event $y_i \notin (-c_{\pi(i)}, c_{\pi(i)})$. Let $1 < m \leq n$ and let s_i be a set of size $(m-1 \wedge i-1)$ of integers $1,2,\ldots,i-1$. Then

$$\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c_i, c_i)\}$$

$$\leq \Pr\{A_1\} + \sum_{i=2}^{n} \Pr\{A_i\} - \Pr\{A_i \cap [\bigcup_{j \varepsilon s_i} A_j]\}$$

$$= \Pr\{A_1\} + \sum_{i=2}^{n} \Pr\{A_i \cap [\bigcup_{j \varepsilon s_i} A_j]^c\} .$$

Proof.

$$\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c_i, c_i)\} = \Pr\{\bigcup_{i=1}^{n} y_i \notin (-c_{\pi(i)}, c_{\pi(i)})\}$$

$$= \Pr\{\bigcup_{i=1}^{n} (A_i)\} = \Pr\{A_1\} + \sum_{i=2}^{n} \Pr\{A_i \cap [\bigcup_{j=1}^{i-1} A_j]^c\}$$

$$\leq \Pr\{A_1\} + \sum_{i=2}^{n} \Pr\{A_i \cap [\bigcup_{j \varepsilon s_i} A_j]^c\} = \Pr\{A_1\} + \sum_{i=2}^{n} \Pr\{A_i\} - \Pr\{A_i \cap [\bigcup_{j \varepsilon s_i} A_j]\} .$$

Remarks.

1) To implement this theorem requires being able to determine $\Pr\{A_i \cap [\bigcup_{j \varepsilon s_i} A_j]^c\}$, which requires being able to integrate over m dimensional marginal distributions.

2) For $m = 2$ the above theorem is equivalent to Hunter's Theorem which is: Let $\underset{\sim}{x}$ be $\sim N(0,\Sigma)$. Let T be any maximal tree connecting the x_i and forming no circuit and e_{ij} be edges on this tree then

$$\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c_i,c_i)\} \le$$

$$\sum \Pr\{x_i \notin (-c_i,c_i)\} - \sum_{e_{ij} \in T} \Pr\{x_i \notin (-c_i,c_i) \cap x_j \notin (-c_j,c_j)\}.$$

3) This theorem still leaves unsolved how to choose the "best" permutation $\pi(\cdot)$ of x_i and sets s_j to give the lowest upper bound for $\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c_i,c_i)\}$. At least one "best" permutation set s_i exists since there are a finite number of permutations and sets s_i. For the case $m = 2$ this is equivalent to finding the best tree T for the theorem mentioned in remark 2. This is possible to do; the procedure is given by Theorem AP-1 in the Appendix.

4) This approach is distribution free. The theorem is true for all unions of any collection of events A_i: $i = 1,\ldots,n$. The fact that A_i was $y_i \notin (-c_{\pi(i)},c_{\pi(i)})$ where $y_i \sim N(0,\Sigma)$ and $\pi(\cdot)$ is a permutation of $1,\ldots,n$; was not used in the proof. The next two approaches will be less general.

<u>Corollary A-IC</u>. Group Subtraction Overlap.

Let s_1 and s_2 be sets of integers from $\{1,\ldots,n\}$.

Let ξ_1 be the upper bound for $\Pr\{\bigcup_{i \in s_1} A_i\}$.

Let ξ_2 be the upper bound for $\Pr\{\bigcup_{i \in s_2} A_i\}$.

Let s_3 be the intersection of the sets s_1 and s_2.

Then

$$\Pr\{\bigcup_{i\varepsilon(s_1\cup s_2)} A_i\} \le \Phi_1 + \Phi_2 - \Pr\{\bigcup_{i\varepsilon s_3} A_i\} .$$

Proof.

$$\Pr\{\bigcup_{i\varepsilon(s_1\cup s_2)} A_i\} = \Pr\{\bigcup_{i\varepsilon s_1} A_i\} + \Pr\{\bigcup_{i\varepsilon s_2} A_j\} - \Pr\{(\bigcup_{i\varepsilon s_1} A_i)\cap(\bigcup_{i\varepsilon s_2} A_j)\}$$

$$\le \Pr\{\bigcup_{i\varepsilon s_1} A_i\} + \Pr\{\bigcup_{i\varepsilon s_2} A_j\} - \Pr\{\bigcup_{i\varepsilon(s_1\cap s_2)} A_i\}$$

$$\le \Phi_1 + \Phi_2 - \Pr\{\bigcup_{i\varepsilon s_3} A_i\} .$$

Approach II. Conditional Multiplicative Approach.

The following theorem is a corollary of a theorem from Glaz and Johnson (1984).

Theorem A-II. Let $\underset{\sim}{x}$ be an $n \times 1$ vector distributed $N(0,\Sigma)$ where all off diagonal elements of $-D\Sigma^{-1}D$ are nonnegative for some D, where D is a diagonal matrix with elements ± 1. Let $\underset{\sim}{y}$ be any permutation $\pi(\cdot)$ of the elements of $\underset{\sim}{x}$ and A_i^c be the event $y_i \varepsilon(-c_{\pi(i)}, c_{\pi(i)})$. Let $1 < m \le n$ and let s_i be a set of size $(m-1 \wedge i-1)$ of integers $i,2,\ldots,i-1$ for $i = 2,\ldots,n$.

Then

$$\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c_i, c_i)\} \le 1 - \prod_{i=1}^{n} \Pr\{A_i^c \mid \bigcap_{j\varepsilon s_i} A_j^c\} .$$

Proof. Part One.

$$\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c_i, c_i)\} = 1 - \Pr\{\bigcap_{i=1}^{n} (x_i \varepsilon (-c_i, c_i))\} = 1 - \Pr\{\bigcap_{i=1}^{n} (y_i \varepsilon (-c_{\pi(i)}, c_{\pi(i)}))\}$$

$$= 1 - \prod_{i=1}^{n} \Pr\{y_i \varepsilon (-c_{\pi(i)}, c_{\pi(i)}) | y_j \varepsilon (-c_{\Pi(j)}, c_{\Pi(j)}),\ j=1,\ldots,i-1\} .$$

Part Two. By a theorem in Karlin and Rinott (1980)

(a) $|\underset{\sim}{y}|$ has an MTP_2* distribution when $\underset{\sim}{x} \sim N(0,\Sigma)$ and the elements of $-D\Sigma^{-1}D$ are nonnegative where D is any diagonal matrix with elements ± 1 .

(b) $y_i \varepsilon (-c_{\pi(i)}, c_{\pi(i)})$ for all $i \Longleftrightarrow |y_i| \varepsilon [0, c_{\pi(i)})$ for all i and these are monotone sets of the same type.

It hence follows from (a), (b) and theorem 2.3 part (1) in Glaz and Johnson (1984) that

$$\Pr\{y_i \varepsilon (-c_{\pi(i)} | y_j \varepsilon (-c_{\pi(j)}, c_{\pi(j)}) \text{ for all } j \varepsilon s_i\} \leq$$

$$\Pr\{y_i \varepsilon (-c_{\pi(i)}, c_{\pi(i)}) | y_j \varepsilon (-c_{\pi(j)}, c_{\pi(j)}),\ j=1,\ldots,i-1\}$$

which implies that

$$\prod_{i=1}^{n} \Pr\{y_i \varepsilon (-c_{\pi(i)}, c_{\pi(i)}) | y_j \varepsilon (-c_{\pi(j)}, c_{\pi(j)}) \text{ for all } j \varepsilon s_i\} \leq$$

$$\prod_{i=1}^{n} \Pr\{y_i \varepsilon (-c_{\pi(i)}, c_{\pi(i)}) | y_j \varepsilon (-c_{\pi(j)}, c_{\pi(j)}),\ j=1,\ldots,i-1\}$$

*See Karlin and Rinott 1980 for a definition of MTP_2.

which when applied to the result of part one implies

$$\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c_i, c_i)\} \leq 1 - \prod_{i=1}^{n} \Pr\{y_i \varepsilon (-c_{\pi(i)}, c_{\pi(i)}) | y_j \varepsilon (-c_{\pi(j)}, c_{\pi(j)}),\ j \varepsilon s_i\}$$

$$= 1 - \prod_{i=1}^{n} \Pr\{A_i^c | \bigcup_{j \varepsilon s_i} A_j^c\} .$$

Remarks.

1) To implement this theorem requires being able to determine $\Pr\{A_i^c | \bigcup_{j \varepsilon s_i} A_j^c\}$ which requires being able to integrate over m dimensional marginal distributions.

2) For $m=2$ the above theorem is equivalent to the following:

Let $x \sim N(0,\Sigma)$ where the off diagonal elements of $-D\Sigma D$ are nonnegative when D is some diagonal matrix with elements ± 1. Let T be any maximal tree connecting the x_i and forming no circuit WLOG let x_1 be a terminal mode of this tree and let all connections be directed toward x_1. Let e_{ij} be the directed edge connecting x_j to x_i i.e. $(x_j \xrightarrow{e_{ij}} x_i)$ and let A_i^c be the event $x_i \varepsilon (-c_i, c_i)$. Then

$$\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c_i, c_i)\} \geq 1 - [\Pr\{A_1\}^c \prod_{e_{ij} \varepsilon T} \Pr\{A_j^c | A_i^c\}].$$

3) As in Theorem A-I for m=2 the best tree T can be found using the procedure of theorem AP-1 in the Appendix. For $m \geq 3$ it is harder to find the "best" permutation $\pi(\cdot)$ and sets s_i to get the lowest possible upper bounds from this method.

4) This theorem is not distribution free it depends highly on the fact that $|x|$ is MTP_2 and that intervals of the form $[0, c_i)$ are monotone of the same type.

Approach III. Independent Subunit Conservative Bounds.

Theorem A-III. (Khatri 1970. Proved by Bechner 1987). Given $\underset{\sim}{x} \sim N(0,\Sigma)$ and $\{W_1\},\{W_2\},\ldots,\{W_F\}$ are disjoint collections of x_i such that

$$\bigcup_{f=1}^{F} \{W_f\} = \{x_1,x_2,\ldots,x_n\} .$$

Then

$$\Pr\{\bigcap_{i=1}^{n} x_i \varepsilon (-c_i,c_i)\} \geq \prod_{f=1}^{F} \Pr\{\bigcap_{x_i \varepsilon W_F} x_i \varepsilon (-c_i,c_i)\}$$

which implies

$$\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c_i,c_i)\} \leq 1 - \prod_{f=1}^{F} \Pr\{\bigcap_{x_i \varepsilon W_F} x_i \varepsilon (-c_i,c_i)\} .$$

Proof. The proof in Khatri's article was incorrect. It has come to my attention that the theorem has been subsequently proven by Bill Bechner at the University of Texas Austin but the proof has not been published yet.

Remarks.

1) By itself, this theorem is not any superior to the Dunn-Sidak method but, as we shall see later, is useful when combined with the two previous approaches mentioned in this chapter.

2) This theorem is highly nondistribution free and can be used only on convex symetric probability regions centered at the origin (or compliments of such regions) involving densities where $f(x) = f(-x)$.

Corollary A-IIIC. Given $\underset{\sim}{x} \sim N(0,\Sigma)$ and $\{W_1\},\{W_2\},\ldots,\{W_F\}$ are disjoint collections of the x_i such that

$$\bigcup_{f=1}^{F} \{W_f\} = \{x_1,x_2,\ldots,x_n\}$$

and ϕ_f is an upper bound for $\Pr\{\bigcup_{i \in W_f} x_i \notin (-c_i, c_i)\}$. Then

$$\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c_i, c_i)\} \leq 1 - \prod_{f=1}^{F} (1-\phi_f) .$$

<u>Proof</u>.

$$\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c_i, c_i)\} \overset{\text{By Thm. A-III}}{\leq} 1 - \prod_{f=1}^{F} \Pr\{\bigcap_{i \in W_f} x_i \in (-c_i, c_i)\}$$

$$= 1 - \prod_{f=1}^{F} \left[1 - \Pr\{\bigcup_{i \in W_f} x_i \notin (-c_i, c_i)\}\right]$$

$$\leq 1 - \prod_{f=1}^{F} (1-\phi_f) .$$

by repeated application
of the distributive
multiplication inequality

III. COMPARING THE METHODS.

Theorems A-I and A-II have been presented in a way to enable us to use m dimensional marginals for $m \geq 2$. To keep the statements and proofs of the theorems simple for the remainder of this technical report m = 2 will be assumed whenever theorems A-I and A-III are used. All of the results in this chapter could be extended to comparable results for $m \geq 3$.

Since there are several different methods to use, the question arises as to under which conditions is one method better than the others (i.e., gives closer bounds). This question is explored further in the present chapter.

Thereom C-1. Let $\underset{\sim}{x}$ and A_i be as described in Theorems A-I, A-II and A-III. Let n = 2 then (i) the intersection subtraction method (Theorem A-I) upper bound is exactly the same as the conditional multiplicative method (Theorem A-II) upper bound and is exact. (ii) When the correlation is non zero, the Dunn-Sidak upper bound result is inferior to the upper bound of theorems A-I and A-II.

Proof. (i) The A-I upper bound for $\Pr\{A_1 \cup A_2\}$ is $\Pr\{A_1\}+\Pr\{A_2\}-\Pr\{A_1 \cap A_2\}$ which equals $\Pr\{A_1 \cup A_2\}$ by elementary axioms of probability. The A-II upper bound for $\Pr\{A_1 \cup A_2\}$ is

$$1-\Pr\{A_1^c\} \Pr\{A_2^c | A_1^c\} = 1-\Pr\{A_1^c \cap A_2^c\} = \Pr\{A_1 \cup A_2\}$$

(ii) Theorem 1 in Sidak (1967) and its proof show that $\Pr\{A_1 \cup A_2\}$ is a monotonically decreasing function of the absolute value of the correlation coefficient for fixed variances σ_1^2 and σ_2^2 where again $A_i = \Pr\{x_i \notin (-c_i, c_i)\}$.

Theorem C-2. Let $\underset{\sim}{x}$ and A_i, $i=1,\ldots,n$ be as described in theorems A-I, A-II and A-III. For $n > 2$, assuming $m = 2$ and that the conditions are met for the conditional multiplicative (A-II) upper bound, then for any upper bound given by either overlap subtraction (A-I) or by Dunn-Sidak (I-1) there exists an equal (or superior if Σ is not diagonal) upper bound given by (A-II).

Proof. First. For a given directed tree T with edges e_{ij} and the root at x_1 as described in remark 2 of theorem (A-II), the (A-II) upper bound is

$$1 - [\Pr(A_1^c) \prod_{e_{ij} \in T} (A_i^c | A_j^c)] = 1 - \Pr(A_1^c) \prod_{e_{ij} \in T} \frac{\Pr(A_i^c \cap A_j^c)}{\Pr(A_j^c)}$$

$$\leq 1 - \Pr(A_1)^c \prod_{e_{ij} \in T} \frac{\Pr(A_i^c)\Pr(A_j^c)}{\Pr(A_j^c)} = 1 - \prod_{i=1}^{n} \Pr(A_i)^c$$

by theorem
1 in Sidak
(The inequality is strict if Σ is not diagonal)

= Dunn-Sidak upper bound .

Second. For a given directed tree T with edges e_{ij} and the root at x_1 as described in remark 2 of theorems (A-II) and (A-I). Let T_2 be the tree connecting the first two elements in T and T_i be the sub tree of T connecting the first i elements in T $i=3,\ldots,n$. Let $(\text{A-I})_i$ and $(\text{A-II})_i$ be the A-I and A-II upper bounds for $\Pr \bigcup_{j \in T_i} (A_j)$. Then by theorem C-I we have

$$(\text{A-I})_2 = (\text{A-II})_2 \ .$$

Now

Where x_3 is the third element of the tree and j_{T_2} is the item in T_2 to which x_3 is connected.

$$(A\text{-}I)_3 = (A\text{-}I)_2 + Pr(A_{x_3}) - Pr(A_{x_3} \cap A_{j_{T_2}})$$

$$= (A\text{-}I)_2 + Pr(A_{x_3} \cap A^c_{j_{T_2}})$$

$$= 1 - [1-(A\text{-}I)_2 - Pr(A_{x_3} \cap A^c_{j_{T_2}})]$$

$$\geq 1 - [[1-(A\text{-}I)_2] \cdot \frac{Pr(A^c_{j_{T_2}}) - Pr(A_{x_3} \cap A^c_{j_{T_2}})}{Pr(A^c_{j_{T_2}})}]$$

Since $(1-(A\text{-}I)_2) \leq (A_{j_{T_2}})^c$

$$= 1 - [[1-(A\text{-}I)_2] \cdot \frac{Pr(A^c_{j_{T_2}} \cap A^c_{x_3})}{Pr(A^c_{j_{T_2}})}]$$

$$\geq 1 - [[1-(A\text{-}II)_2] \cdot Pr(A^c_{x_3} \mid A^c_{j_{T_2}})]$$

Since $(A\text{-}I)_2 \geq (A\text{-}II)_2$

$$= (A\text{-}II)_3 .$$

Thus $\qquad (A\text{-}I)_3 \geq (A\text{-}II)_3$.

Following this proof inductively expressing $(A\text{-}I)_4$ and $(A\text{-}II)_4$ in terms of $(A\text{-}I)_3$ and $(A\text{-}II)_3$ and x_4 the fourth element of tree T_4 we will get $(A\text{-}I)_4 \geq (A\text{-}II)_4$ etc. until obtaining $(A\text{-}I)_n \geq (A\text{-}II)_n$ for any $n > 2$. But $T_n = T$ and thus the upper bound from using A-I on T is higher than the upper bound from using A-II on T which completes the proof.

Theorem C-3. Let T be as described in remark 2 of theorems A-II and A-III and let σ_i^2 and c_i be constant. If $\max_T |P_{ij}|$ is bounded by some value less than one then as $n \to \infty$, the Sidak-Dunn assumption of independence eventually produces a better upper bound for

$$\Pr\{\bigcup_{i=1}^{n} A_i\}$$

than does the intersection subtraction method (A-I) and in fact for n large enough, (A-I) will produce an upper bound for

$$\Pr\{\bigcup_{i=1}^{n} A_i\}$$

which is larger than 1.

Proof. The intersection subtraction method upper bound is

$$\sum \Pr\{A_i\} - \sum_{e_{ij} \in T} \Pr\{A_i \cap A_j\} =$$

$$\Pr\{A_1\} + \sum_{e_{ij} \in T} [\Pr\{A_i\} - \Pr\{A_i \cap A_j\}] \overset{*}{\geq}$$

*By Theorem 1 of Sidak since $|P_{ij}|$ is bounded below 1, $\Pr\{A_i\} - \Pr\{A_i \cap A_j\} \geq \eta > 0$ for all i and j. See footnote below.

$$\Pr\{A_1\} + \sum_{e_{ij} \in T} [\underset{\text{for some } \eta > 0}{\eta}] = \Pr\{A_i\} + (n-1)\eta \underset{\text{for n large enough}}{>} 1 > 1 - \prod_{1}^{n} (1 - \Pr\{A_i\}) = \text{independence upper bound}$$

*Since c_i and σ_i^2 are constant and $\max_T |P_{ij}|$ is bounded by some value less than one, it follows that $\Pr\{A_i\} = k_1$, for k_1 some constant and by Theorem 1 in Sidak that $\Pr\{A_i \cap A_j\} \leq k_2$ for $k_2 < k_1$. Therefore, $[\Pr\{A_i\} - \Pr\{A_i \cap A_j\}] \geq \eta$ for some constant $\eta > 0$.

Conclusion.

Theorem C-2 shows that whenever it applies, the conditional multiplicative (Theorem A-II) always gives tighter upper bounds than do either the intersection subtraction (Theorem A-I) or the Sidak-Dunn (Thereom I-1) methods. Hence (A-II) should be used in preference to (A-I) and (I-1) to obtain upper bounds whenever the data is jointly MTP_2. Theorem C-1 implies that for small n, the intersection subtraction method (A-I) gives tighter upper bounds than does the Dunn-Sidak method (I-1). Theorem C-3 implies that as n gets larger, the Dunn-Sidak (I-1) upper bounds tend to become superior to those of the intersection subtraction method (A-I) for cases with the same c_i, and $|P_{ij}|$ bounded away from one. The Dunn-Sidak method has been compared with the intersection subtraction method in more detail (see Hoover (1986)).

IV. COMBINING THE APPROACHES.

The three methods A-I, A-II and A-III can be combined together to give lower upper bounds to $\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c_i, c_i)\}$ than could be obtained by using any of these methods, or the Dunn-Sidak Theorem alone. In fact, whenever Σ is nondiagonal, there will exist some combination of A-I, A-II and A-III which gives a lower upper bound to $\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c_i, c_i)\}$ than does the currently used Dunn-Sidak method. It would be nice to be able to, for a given situation, determine the best combination of A-I, A-II and A-III in terms of giving the lowest upper bounds for $\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c_i, c_i)\}$. Unfortunately in most cases this would be computationally impractical to do. Nevertheless it is possible to develop procedures which combine the methods in ways that produce good results.

<u>Example M-one.</u>

At this point, a description and example of how the three methods can be used together should be given. Let us assume $\underset{\sim}{x}$ is an 8×1 vector with a $N(0,\Sigma)$ distribution and that a directed tree with the following links has been set up

$$\begin{array}{ccccccccc} & & & & x_4 & & & & \\ & & & & \downarrow & & & & \\ x_1 & \leftarrow & x_2 & \leftarrow & x_3 & \leftarrow & x_7 & \leftarrow & x_8 \\ & & & & \uparrow & & & & \\ & & & & x_5 & & & & \\ & & & & \uparrow & & & & \\ & & & & x_6 & & & & \end{array}$$

We now define three types of links

S - Intersection subtraction (A-I)

M - Conditional multiplicative (A-II)

I - Conservative independent subunit (A-III)

These links correspond to which method is being used. We can now assign these link types to such links on the tree for instance

$$\begin{tikzcd}
 & & x_4 \arrow[d, "S"] & & \\
x_1 & x_2 \arrow[l, "S"'] & x_3 \arrow[l, "I"'] & x_7 \arrow[l, "M"'] & x_8 \arrow[l, "S"'] \\
 & & x_5 \arrow[u, "M"'] & & \\
 & & x_6 \arrow[u, "M"'] & &
\end{tikzcd}$$

The first step we wish to take is to identify the continuous sets of variables connected by M links and calculate the upper bounds for $\Pr\{\bigcup_{set} A_i\}$ where A_i is as defined in Theorem A-II. In the above example there is only one such set $x_6 \overset{M}{\rightarrow} x_5 \overset{M}{\rightarrow} x_3 \overset{M}{\leftarrow} x_7$. The (A-II) upper bound for

$$\Pr\{\bigcup_{i=3,5,6,7} (A_i)\} \text{ is } \quad {}^{*}1-[\Pr\{A_3^c\}\cdot\Pr\{A_7^c|A_3^c\}\cdot\Pr\{A_5^c|A_3^c\}\cdot\Pr\{A_6^c|A_5^c\}]$$

by repeated applications of A-II.

Note the order that (A-II) was applied does not matter. For instance if we took $x_6 \overset{M}{\rightarrow} x_5 \overset{M}{\rightarrow} x_3 \overset{M}{\rightarrow} x_7$ to get

$$\Pr\{\bigcup_{i=3,5,6,7} (A_i)\} \le \quad {}^{**}1-[\Pr\{A_7^c\}\cdot\Pr\{A_3^c|A_7^c\}\cdot\Pr\{A_5^c|A_3^c\}\cdot\Pr\{A_6^c|A_5^c\}],$$

the reader can check that ** gives the same value as *.
So let ‡ be the (A-II) upper bound for $\Pr\{\bigcup_{i=3,5,6,7} A_i\}$.

The next step is to use Theorem (A-I) to compute upper bounds using those calculated in the first step and incorporating the S links. In the above example this will give us

$$\Pr\{\bigcup_{i=1,2}(A_i)\} \leq \Pr\{A_1\} + \Pr\{A_2\} - \Pr\{A_1 \cap A_2\} \text{ by theorem (A-I)}$$

and

$$\Pr\{\bigcup_{i=3,5,6,7,4,8}(A_i)\} \leq \Phi + \Pr\{A_4\} + \Pr\{A_8\} - \Pr\{A_4 \cap A_3\} - \Pr\{A_8 \cap A_7\} \text{ by a}$$

slight modification of Theorem (A-I) .

Let Φ_1 be the upper found for $\Pr\{\bigcup_{i=1,2}(A_i)\}$ which in this case was $\Pr\{A_1\}+\Pr\{A_2\}-\Pr\{A_1 \cap A_2\}$ and let Φ_2 be the upper bound for the $\Pr\{\bigcup_{i=3,5,6,7,4,8}(A_i)\}$ which was derived above.

The last step is to incorporate the I link between x_3 and x_2 by using corollary (A-IIIC) to give us the upper bound for

$$\Pr\{\bigcup_{i=1}^{8}(A_i)\} \leq 1 - (1-\Phi_1) \quad (1-\Phi_2) \ .$$

It should be noted that for real cases where a tree has been determined, assignment of S and I links is arbitrary, whereas M links can be assigned only if the joint distribution of the entire group connected by M links meets the MTP_2 conditions of theorem (A-II).

It should also be mentioned that the M links were handled first in step 1 because theorem (A-II) can only be used with M links. There is no way to incorporate other types of links into the results. The S links were then incorporated before the I links were since doing so gives superior results (lower upper bounds).

Example M-two.

This is another, different approach which could be used to combine the methods together to obtain an upper bound. Let us assume $\underset{\sim}{x}$ is a 4×1 vector with a $N(0,\Sigma)$ distribution and that a directed tree with the following links has been set up

$$x_1 \leftarrow x_2 \leftarrow x_3 \leftarrow x_4 \quad .$$

Let $\Pr\{A_i\}$ have the usual definition for $i = 1,\ldots,4$ and assume that the $(|x_1|,|x_2|,|x_3|)$ are jointly MTP_2 and the $(|x_2|,|x_3|,|x_4|)$ are jointly MTP_2 but $(|x_1|,|x_2|,|x_3|,|x_4|)$ is not jointly MTP_2. Then by Theorem (A-II) an upper bound for $\Pr\{A_1 \cup A_2 \cup A_3\}$ is

$$\phi_1 = 1-[\Pr\{A_1^c\} \cdot \Pr\{A_2^c \cap A_1^c\}/\Pr\{A_1^c\} \cdot \Pr\{A_2^c \cap A_3^c\}/\Pr\{A_2^c\}] \quad *$$

and by Theorem (A-II) an upper bound for $\Pr\{A_2 \cup A_3 \cup A_4\}$ is

$$\phi_1 = 1-[\Pr\{A_2^c\} \cdot \Pr\{A_3^c \cap A_2^c\} \; /\Pr\{A_2^c\} \cdot \Pr\{A_4^c \cap A_3^c\}/\Pr\{A_3^c\} \quad .$$

Using ϕ_1 and ϕ_2 along with corollary A-IC and

$$S_1 = A_1, A_2, A_3; \; S_2 = A_2, A_3, A_4; \; S_3 = A_2, A_3;$$
$$S_1 \cup S_2 = A_1, A_2, A_3, A_4$$

gives:

$$\Pr\{\bigcup_{i=1}^{4} A_i\} \le \phi_1 + \phi_2 - \Pr\{A_2 \cup A_3\}$$

* Note $\Pr\{A_i^c | A_j^c\} = \dfrac{\Pr\{A_i^c \cap A_j^c\}}{\Pr\{A_j^c\}}$

It would be nice if it were possible to find the best tree T and method combination set up which gave the lowest upper bounds. To do so would require a quadratic program which, except for small n would be computationally impractical to solve.

It might be possible to develop a "good" procedure to combine the three approaches in such a way to give a lower upper bound than does the Dunn-Sidak method when Σ is not diagonal and hopefully give an upper bound which is close to the lowest possible upper bound achievable from any procedure combining the methods.

The author believes that there is one particular situation where one can find the lowest upper bound for $\Pr\{\bigcup_{i=1}^{n} A_i\}$ possible from combining the three procedures and show that this method is indeed the lowest possible upper bound. That situation is when the absolute values of all the elements in $\underset{\sim}{x}$ (i.e. $|x_1|, |x_2|, \ldots, |x_n|$) are jointly MTP_2.

<u>Hypothesis M-1</u>. Let $\underset{\sim}{x} \sim N(0,\Sigma)$ if $|\underset{\sim}{x}|$ is also jointly MTP_2 i.e. for some diagonal matrix D with elements ± 1 such that the off diagonal elements of $-D\Sigma D$ are all nonnegative, then the upper bound for $\Pr\{\bigcup_{i=1}^{n} A_i\}$ or $\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c_i, c_i)\}$ given by (A-II) from using any tree T is lower than the upper bound for $\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c_i, c_i)\}$ given from using any combination of A-I, A-IC A-I, A-III and A-IIIC on that same tree T.

The proof would be quite complicated since there are so many ways to combine the different methods together which must be considered.

Given that hypothesis M-1 is true, then to get the lowest possible upper bound for $\Pr\{\bigcup_{i=1}^{n} A_i\}$ from any combination of the three approaches, we use the (A-II) approach entirely on T' the best tree derived in remark 2 of theorem (A-II) and theorem (AP-1) in the Appendix.

V. EXAMPLES OF APPLYING THE COMBINED METHODS.

Example One. Consider the MA(5) model. Let x_i be iid N(0,1) and let $y_j = \sum_{i=1}^{5} x_{j+i}$. Then $P_{y_j,y'_j} = \max(\frac{5-|j-j'|}{5},0)$ and $\sigma_{y_j} = \sqrt{5}$.

Let the set $R = \{y_1,y_2,y_3,y_7,y_8,y_9\}$ and suppose the researcher wants to know upper bounds for

A) $\Pr\{\bigcup_{j\varepsilon R} |y_j| > 1.96\sqrt{5}\}$

B) $\Pr\{\bigcup_{j\varepsilon R} |y_j| > 2.25\sqrt{5}\}$

C) $\Pr\{\bigcup_{j\varepsilon R} |y_j| > 2.50\sqrt{5}\}$

D) A lower bound for c such that $\Pr\{\bigcup_{j\varepsilon R} |y_i| > c\} \leq .10$.

Clearly the best tree T' is $y_1-y_2-y_3-y_7-y_8-y_9$ and $p_{12} = p_{23} = p_{78} = p_{89} = .80$ while $p_{37} = .20$. It can be shown that for any continuous group of size 3 or more on T' that there is no MTP_2 density so we are dealing only with S and I links. Below answers to A, B and C are shown using the procedure of example M-one in chapter IV with

(i) all links as I links

(ii) all links as S links

(iii) $y_1-y_2,y_2-y_3,y_7-y_8,y_8-y_9$ as S links y_3-y_7 as an I link

	METHOD		
	(i) All I links	(ii) All S links	(iii) y_3-y_7 I link all other S links
A) $\Pr\{\bigcup_{j\varepsilon R} \|y_j\|>1.96\sqrt{5}\}$	.2649	.2051	.2010
B) $\Pr\{\bigcup_{j\varepsilon R} \|y_j\|>2.25\sqrt{5}\}$	.1380	.1076	.1057
C) $\Pr\{\bigcup_{j\varepsilon R} \|y_j\|>2.50\sqrt{5}\}$	.0722	.0569	.0564

The combined method (iii) works better than does either the pure I method or the pure S method although the improvements from (iii) over (ii) are not spectacular. The upper bound ratio of method (iii) over method (ii) ranged from .991 for C) to .980 for A).

The answer to D using methods (i), (ii) and (iii) respectively is

(i) 2.383

(ii) 2.283

(iii) 2.275

<u>Example Two</u>. Let $(x_1,\ldots,x_6) \sim N(0,\Sigma)$ where

$$\Sigma = \begin{matrix} 1 & & & & & \\ .9 & 1 & & & & \\ .9 & .9 & 1 & & & \\ .9 & .9 & .9 & 1 & & \\ .5 & .9 & .9 & .9 & 1 & \\ .3 & .5 & .9 & .9 & .9 & 1 \end{matrix}$$

Let the researcher be interested in upper bounds for

A) $\Pr\{\bigcup_{i=1}^{6} |x_i| > 1.96\}$

B) $\Pr\{\bigcup_{i=1}^{6} |x_i| > 2.25\}$

C) $\Pr\{\bigcup_{i=1}^{6} |x_i| > 2.50\}$

and

D) A lower bound for c such that $\Pr\{\bigcup_{i=1}^{6} |x_i| > c\} \le .05$.

Clearly a best tree T' is

$$x_1 - x_2 - x_3 - x_4 - x_5 - x_6$$

It turns out that while x itself is not jointly MTP_2 the sub-vectors $(|x_1|, |x_2|, |x_3|, |x_4|)$ and $(|x_3|, |x_4|, |x_5|, |x_6|)$ are each jointly MTP_2. So it is possible to use an algorithm similar to that of example M-Two in chapter IV to get an upper bound. We shall call this method (*).

(*) First use (A-II) to get upper bounds for

$$\beta_1 = \Pr\{\bigcup_{i=1}^{4} |x_i| > c\} \quad \text{and} \quad \beta_2 = \Pr\{\bigcup_{i=3}^{6} |x_i| > c\} .$$

Then use (A-I-C) to give

$$\Pr\{\bigcup_{i=1}^{6} |x_i| > c\} \le \beta_1 + \beta_2 - \Pr\{\bigcup_{i=3}^{4} |x_i| > c\} .$$

Below answers to A, B and C are shown using

(i) All links as I links

(ii) All links as S links

(iii) Method (*)

METHOD

	(i) All I links	(ii) All S links	(iii) *
A) $\Pr\{\bigcup_{i=1}^{6} \|x_i\| > 1.96\}$	.2649	.1518	.1491
B) $\Pr\{\bigcup_{i=1}^{6} \|x_i\| > 2.25\}$	.1380	.0792	.0784
D) $\Pr\{\bigcup_{i=1}^{6} \|x_i\| > 2.50\}$	.0722	.0423	.0421

Again the improvement of method (iii) over method (ii) was not spectacular. The upper bound ratio of (iii) over (ii) was .995 in C, .990 in case B and .982 in case A.

The answer to D using methods (i), (ii) and (iii) respectively is

(i) 2.635

(ii) 2.436

(iii) 2.433

Example Three. This is based on problem E on page 230 of Drapper and Smith. A linear regression is developed to predict corn yield per acre in Iowa. Five of the coefficients in the final equation are 1. Year, 2. July Rain, 3. July Temperature, 4. August Rain and 5. August Temperature. Ignoring the constant coefficient b0; call the five coefficients, b1, b2, b3, b4, b5, and the estimated coefficients $\hat{b}1$, $\hat{b}2$, $\hat{b}3$, $\hat{b}4$ and $\hat{b}5$ respectively. The correlation matrix of $\hat{\underset{\sim}{b}}$ is

[illegible] =

$$\begin{matrix} 1 & & & & \\ -.1160 & 1 & & & \\ .4870 & .3445 & 1 & & \\ -.1725 & .1215 & -.0914 & 1 & \\ -.2076 & .3362 & -.3042 & .3339 & 1 \end{matrix}$$

The best tree T turns out to be

$$\hat{b}_1 - \hat{b}_3 \quad \hat{b}_2 - \hat{b}_5 - \hat{b}_4 \ .$$

The distribution of $\underset{\sim}{\hat{b}}$ is not MTP_2 but the distribution of $\hat{b}_2$, $\hat{b}_4$ and $\hat{b}_5$ is MTP_2 however.

Let us assume that the researcher wants to have upper bounds for

A) $\Pr\{\bigcup_{i=1}^{5} b_i \notin \hat{b}_i \pm 1.96\,\sigma_{\hat{b}_i}\}$

B) $\Pr\{\bigcup_{i=1}^{5} b_i \notin \hat{b}_i \pm 2.25\,\sigma_{\hat{b}_i}\}$

C) $\Pr\{\bigcup_{i=1}^{5} b_i \notin \hat{b}_i \pm 2.50\,\sigma_{\hat{b}_i}\}$

and

D) A lower bound for c such that $\Pr\{\bigcup_{i=1}^{5} b_i \notin \hat{b}_i \pm c\sigma_{\hat{b}_1}\} \le .10$

and also the researcher wants a lower bound for c so that $\Pr\{\bigcup_{i=1}^{5} \hat{b}_i > (b_i + c\sigma_{\hat{b}_i})\} \le .10$. We can calculate these bounds by using the procedure in example M-one section IV with (i) all S links, (ii) all I links or (iii) the following links $\hat{b}_1 \overset{S}{—} \hat{b}_3 \overset{I}{—} \hat{b}_2 \overset{M}{—} \hat{b}_5 \overset{M}{—} \hat{b}_4$.

The upper bounds for A, B and C are

	METHOD (i) All I links	METHOD (ii) All S links	METHOD (iii)
A) $\Pr\{\bigcup_{i=1}^{5} b_i \notin \hat{b}_i \pm 1.96\,\sigma_{\hat{b}_i}\}$	.2262	.2245	.2156
B) $\Pr\{\bigcup_{i=1}^{5} b_i \notin \hat{b}_i \pm 2.25\,\sigma_{\hat{b}_i}\}$	.1164	.1135	.1117
C) $\Pr\{\bigcup_{i=1}^{5} b_i \notin \hat{b}_i \pm 2.50\,\sigma_{\hat{b}_i}\}$	.0605	.0589	.0586

The upper bounds for D using (i), (ii) and (iii) respectively are

(i) 2.319

(ii) 2.310

(iii) 2.306

The probability ratio of method (iii) to method (ii) was .995 in case C, .984 in Case B and .960 in case A.

Analysis.

In the three examples of section V, the merged methods produced lower upper bounds for probability than did any of the single methods used alone. In all three examples using only S links was the best competitor being from .960 to .995 as efficient. One difficulty involved with merging the methods is that it is not simple to figure out "how" to set up the combined links. It would appear difficult to come up with a simple universal algorithm to do this meaning that one would have to do this oneself by making educated guesses as to which combination of methods to use.

VI. TABLES OF RESULTS FROM COMBINED (A-I) AND (A-III) METHODS.

At this point, we will try to explore, in some systematic fashion, when combining the methods gives improved results over using any of the methods singly. For simplicity, Method (A-II) the conditional multiplicative method is not being considered. If (A-II) links could be included in the combined methods studied in the next few tables one would expect that even lower upper bounds would be obtained. The methods will be combined together in the fashion of example M-one in Chapter IV.

Tables A, B and C look at various combinations of S and I links compared to the best of using all S links, or using all I links, for calculating upper bounds of $\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c,c)\}$ where $n = 4,6,10$ respectively, under different correlation structures for the links. In all three tables c is set equal to 1.96 and 2.50 while the best tree T' is assumed to be a string of the form $(x_1 - x_2 - \cdots - x_n)$.

In table A, link patterns of the form $x_1 \overset{S}{-} x_2 \overset{I}{-} x_3 \overset{S}{-} x_4$ and $x_1 \overset{S}{-} x_2 \overset{S}{-} x_3 \overset{I}{-} x_4$ are compared to the best of $x_1 \overset{S}{-} x_2 \overset{S}{-} x_3 \overset{S}{-} x_4$ and $x_1 \overset{I}{-} x_2 \overset{I}{-} x_3 \overset{I}{-} x_4$. The combined methods do give better results than does the pure S method or the conservative assumption of independence when $|\rho_{ij}| \leq .3$ for x_i linked to x_j by the I link. The combined methods do better when c = 2.50 having upper bound ratio values as low as .953 compared to pure methods than they do when c = 1.96 having upper bound ratios by factors as low as .977 compared to pure methods.

In table B, link patterns of $x_1 \overset{S}{-} x_2 \overset{S}{-} x_3 \overset{I}{-} x_4 \overset{S}{-} x_5 \overset{S}{-} x_6$ and $x_1 \overset{S}{-} x_2 \overset{S}{-} x_3 \overset{S}{-} x_4 \overset{I}{-} x_5 \overset{S}{-} x_6$ are compared to the best of $x_1 \overset{S}{-} x_2 \overset{S}{-} x_3 \overset{S}{-} x_4 \overset{S}{-} x_5 \overset{S}{-} x_6$

and $x_1 \overset{I}{-} x_2 \overset{I}{-} x_3 \overset{I}{-} x_4 \overset{I}{-} x_5 \overset{I}{-} x_6$. The combined methods give better results than do the pure methods when $|p_{ij}| \leq .4$ for x_i linked to x_j by the I link. The relative upper bound ratios of the combined methods to the pure methods were somewhat lower (better) when $n = 6$ than when $n = 4$. When $n = 6$ the ratio of upper bound combined/best upper bound pure methods were as low as .955 for $c = 1.96$ and as low as .985 for $c = 2.50$.

In table C, the link patterns of :

(i) $x_1 \overset{S}{-} x_2 \overset{I}{-} x_3 \overset{S}{-} x_4 \overset{I}{-} x_5 \overset{S}{-} x_6 \overset{I}{-} x_7 \overset{S}{-} x_8 \overset{I}{-} x_9 \overset{S}{-} x_{10}$

(ii) $x_1 \overset{S}{-} x_2 \overset{S}{-} x_3 \overset{I}{-} x_4 \overset{S}{-} x_4 \overset{I}{-} x_6 \overset{S}{-} x_7 \overset{I}{-} x_8 \overset{S}{-} x_9 \overset{S}{-} x_{10}$

(iii) $x_1 \overset{S}{-} x_2 \overset{S}{-} x_3 \overset{S}{-} x_4 \overset{S}{-} x_5 \overset{I}{-} x_6 \overset{S}{-} x_7 \overset{S}{-} x_8 \overset{S}{-} x_9 \overset{S}{-} x_{10}$

are compared to the best of

$$x_1 \overset{S}{-} x_2 \overset{S}{-} x_3 \overset{S}{-} x_4 \overset{S}{-} x_5 \overset{S}{-} x_6 \overset{S}{-} x_7 \overset{S}{-} x_8 \overset{S}{-} x_9 \overset{S}{-} x_{10} \quad \text{and}$$

$$x_1 \overset{I}{-} x_2 \overset{I}{-} x_3 \overset{I}{-} x_4 \overset{I}{-} x_5 \overset{I}{-} x_6 \overset{I}{-} x_7 \overset{I}{-} x_8 \overset{I}{-} x_9 \overset{I}{-} x_{10}$$

The combined methods give better results than do the pure methods when $|p_{ij}| \leq .8$ for $c = 1.96$ and $|p_{ij}| \leq .6$ for $c = 2.50$. The probability ratio of combined method/best pure method were as low as .928 for $c = 1.96$ and as low as .977 for $c = 2.50$.

Summarizing tables A, B and C the following tendencies were observed:

1. The combined methods were better with respect to the pure methods as n became larger. It is felt that this will continue to be true for larger values of n than were looked at here.

2. For $|P_{ij}|$ not too close to zero where x_i is linked to x_j by an I link in the combined method, the combined methods became better with respect to the pure methods as $|P_{ij}|$ became smaller.

3. The combined methods were better with respect to the pure methods for the smaller value of $c = 1.96$ versus the larger value of $c = 2.50$. It is felt that this will hold in general for all values of c.

The question remains as to whether there are applications where it is worth the extra effort to compute the merged results to get the lower upper bounds than are given by the pure methods.

TABLE A

UPPER BOUNDS FOR $\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c,c)\}$ OBTAINED FOR N=4, WHEN THE BEST TREE IS A STRING x_1-x_2-x_3-x_4 BY USING DIFFERENT COMBINATIONS OF S AND I LINKS

Correlations along the various links			c = 1.96		c = 2.50	
x_1-x_2	x_2-x_3	x_3-x_4	Best of all S and all I	Smallest P_{ij}=I Rest are S^{ij}	Best of all S and all I	Smallest P_{ij}=I Rest are S^{ij}
.9	.5	.9	.1314	.1357	.0355	.0365
.8	.5	.8	.1469	.1501	.0397	.0405
.9	.4	.9	.1340	.1357	.0360	.0365
.8	.4	.8	.1495	.1501	.0402	.0405
.6	.4	.6	.1684	.1675	.0453	.0451
.9	.3	.9	.1358	.1357	.0362	.0365
.8	.3	.8	.1513	.1501	.0405	.0405
.7	.3	.7	.1621	.1600	.0432	.0432
.6	.3	.6	.1703	.1675	.0451	.0451
.9	.9	.2	.1372	.1362	.0365	.0365
.9	.2	.9	.1372	.1357	.0365	.0365
.9	.1	.9	.1379	.1357	.0366	.0365
.8	.1	.9	.1455	.1430	.0387	.0385
.5	.5	.5	.1721	.1731	.0457	.0464
.5	.4	.5	.1747	.1731	.0487	.0464
.4	.4	.4	.1799	.1779	.0487	.0474
.3	.3	.3	.1854	.1811	.0487	.0480
.3	.2	.3	.1854	.1811	.0487	.0480
.2	.2	.2	.1854	.1837	.0487	.0484

Note for N=4 the upper bound for $\Pr\{\bigcup_{i=1}^{4} x_i \notin (-c,c)\}$ given by the conservative independence assumption is

(i) .1854 for c = 1.96 :

(ii) .0487 for c = 2.50 :

TABLE B

UPPER BOUNDS FOR $Pr\{\bigcup_{i=1}^{n} x_i \notin (-c,c)\}$ OBTAINED FOR N=6, WHEN THE BEST TREE IS A STRING $x_1-x_2-x_3-x_4-x_5-x_6$ BY USING DIFFERENT COMBINATIONS OF S AND I LINKS

Correlations along the various links					c = 1.96		c = 2.50	
x_1-x_2	x_2-x_3	x_3-x_4	x_4-x_5	x_5-x_6	Best of All S and all I	Smallest P_{ij}=I Rest are S	Best of All S and all I	Smallest P_{ij}=I Rest are S^{ij}
.9	.9	.5	.9	.9	.1723	.1733	.0475	.0482
.9	.9	.4	.9	.9	.1749	.1733	.0479	.0482
.8	.8	.4	.8	.8	.2057	.2011	.0562	.0563
.9	.7	.4	.7	.9	.2010	.1970	.0549	.0554
.6	.5	.4	.5	.6	.2499	.2401	.0669	.0666
.4	.4	.4	.4	.4	.2665	.2545	.0702	.0698
.8	.7	.3	.7	.8	.2183	.2107	.0594	.0590
.8	.8	.8	.3	.8	.2075	.2019	.0566	.0564
.3	.3	.3	.3	.3	.2649	.2607	.0722	.0711
.3	.3	.1	.3	.3	.2649	.2607	.0722	.0711

Note for N=6 the upper bound for $Pr\{\bigcup_{i=1}^{6} x_i \notin (-c,c)\}$ given by the conservative independence assumption is

(i) .2649 for c = 1.96:

(ii) .0722 for c = 2.50:

TABLE C

UPPER BOUNDS FOR $\Pr\{\bigcup_{i=1}^{n} x_i \notin (-c,c)\}$ OBTAINED FOR N=6 WHEN THE BEST TREE T IS A STRING $x_1-x_2-x_3-x_4-x_5-x_6-x_7-x_8-x_9-x_{10}$ [AND p_{ij} FOR e_{ij} T IS A CONSTANT VALUE P] BY USING DIFFERENT COMBINATIONS OF S AND I LINKS.

n=10

c = 1.96

P	Best of all S and all I	x_5-x_6 I links Rest S links	$x_2-x_3, x_5-x_6, x_7-x_8$ I links Rest S links	$x_2-x_3, x_4-x_5, x_6-x_7, x_8-x_9$ I links Rest S links
.9	.2335	.2458	.2856	.3058
.8	.3029	.2984	.3210	.3341
.7	.3515	.3341	.3450	.3533
.6	.3384	.3606	.3629	.3677
.5	.4012	.3803	.3761	.3783
.4	.4012	.3965	.3871	.3872
.3	.4012	.4012	.3945	.3932
.2	.4012	.4012	.4004	.3979
.1	.4012	.4012	.4012	.3999

c = 2.50

P				
.9	.0663	.0714	.0829	.0887
.8	.0851	.0875	.0946	.0984
.7	.0975	.0979	.1022	.1047
.6	.1060	.1051	.1075	.1090
.5	.1121	.1105	.1112	.1120
.4	.1164	.1138	.1139	.1142
.3	.1175	.1163	.1157	.1157
.2	.1175	.1175	.1169	.1168
.1	.1175	.1175	.1175	.1173

Note for N=10 the upper bound for $\Pr\{\bigcup_{i=1}^{10} x_i \notin (-c,c)\}$ given by the conservative independence assumption is (i) .4012 for c = 1.96: (ii) .1175 for c = 2.50:

VII. TERMINOLOGY

$\underset{\sim}{x}$ - is an $n \times 1$ vector of random variables with a $N(0,\Sigma)$ distribution

P_{ij} - is the correlation between x_i and x_j

A_i - is the event x_i is not in the interval $(-c_i, c_i)$

S_i - is a set of integers representing corresponding numbered events from $A_1, A_2, \ldots, A_{i-1}$ or equivalently corresponding numbered variables from $x_1, x_2, \ldots, x_{i-1}$

$(m \wedge n)$ - is the minimum of m and n

$\Pr\{\cdot\}$ - is the probability that $\{\cdot\}$ occurs

T - is a tree connecting the events $A_1, A_2, \ldots, A_n$ or equivalently the variables $x_1, x_2, \ldots, x_n$

e_{ij} - is a directed link from x_i to x_j or equivalently from A_i to A_j e.g. $x_i \xrightarrow{e_{ij}} x_j$

S link - an S link from x_i to x_j indicates that the upper bound operation which will be performed between these two variables (events) is of the intersection subtraction (Theorem A-I) type.

M link - an M link from x_i to x_j indicates that the upper bound operation which will be performed between these two variables (events) is of the conditional multiplicative (Theorem A-II) type.

I link - an I link from x_i to x_j indicates that the upper bound operation which will be performed between these two variables (events) is of the conservative independent subunit (Theorem A-III) type.

¢ - an upper bound to $\Pr\{\bigcup_{i=1}^{n} A_i\}$.

$A|B$ - $\Pr\{A|B\}$ means the probability of A given B.

VIII. APPENDIX.

Below is presented the theorem to find the "best" tree with which to apply method (A-I) or method (A-II) for the case $m = 2$. Note that a maximal tree is one which has the maximal number of links connecting n elements while forming no loops, ($n-1$ links).

<u>Theorem AP-1</u>. Let $x \sim N(0,\Sigma)$ and let P_{ij} be the correlation coefficient between x_i and x_j. Let T be any maximal tree connecting the individual components of x. Let 0_T be the order statistics of $|P_{ij}|$ for $e_{ij} \varepsilon T$. Set up a tree T' by performing the following process $n-1$ times: Link x_i to x_j for the i,j which maximize $|P_{ij}|$ among all unlinked i,j whose linking does not create a circuit. Then this tree will have the following property $(0_{T'})_j \geq (0_T)_j$ for all other trees T and all $j=1,\ldots,n$.

<u>Proof</u>. The reader is referred to Hoover (1986) or Kruskal (1956).

<u>Remark</u>: Since Sidak's theorem gives that $Pr(A_i \cap A_j)$ is a nondecreasing function of $|P_{ij}|$ the above theorem implies that T' is the best tree on which to apply methods (A-I) and (A-II). Note than when using a combined method for any given situation, T' may not be the best tree, but should still be a good tree to use.

IX. REFERENCES

Beckner, B. (1987). Unpublished document. University of Texas, Austin.

Bolviken, E. (1981). Probability inequalities for the multivariate normal with non-negative partial correlations. Scand. J. Statist., 9, 48-58.

Dunn, O. J. (1958). Estimation of the means of dependent variables. Ann. Math. Stat., 29, 1097-1111.

Glaz, J. and Johnson, B. (1984). Probability inequalities for multivariate distributions with dependence structures. JASA, 79, 436-440.

Hoover, D. (1986). Deriving and applying improved upper bounds for multivariate normal probability outside of N cubes. Stanford University, Department of Statistics Technical Report.

Hoover, D. (1987). Selective inclusion upper and lower bounds. Stanford University, Department of Statistics Ph.D. dissertation.

Hunter, D. (1976). An upper bound for the probability of a union. J. Appl. Prob., 13, 597-603.

Karlin, S. and Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities - I. Multivariate totally positive distributions. J. Mult. Anal., 10, 467-498.

Khatri, G. C. (1970). Further contributions to some inequalities for normal distributions and their applications to simultaneous confidence bounds. Ann. Inst. Math. Statist., 38, 451-458.

Kruskal, J. (1956). On the shortest spanning subtree of a graph and the traveling salesman problem, Proc. Amer. Math. Soc., 7, 48-50.

Sidak, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. JASA, 66, 626-633.

Worsley, K. J. (1982). n improved Bonferoni inequality and applications. Biometrika, 69, 297-302.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 409	2. GOVT ACCESSION NO. [illegible]	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Different Algorithms For Obtaining Upper Bounds To Multivariate Normal Areas Outside Of Origin Centered Rectangles Using Joint Marginal Probabilities		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Donald R. Hoover		8. CONTRACT OR GRANT NUMBER(s) N00014-86-K-0156
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-267
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 1111		12. REPORT DATE September 27, 1988
		13. NUMBER OF PAGES 38
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Bonferonni Bounds, Multivariate Normal, Simultaneous Inference.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

PLEASE SEE FOLLOWING PAGE.

TECHNICAL REPORT NO. 409

20. ABSTRACT

Upper bounds to multivariate normal probability areas outside of n dimensional rectangles centered at the origin are of interest due to their applications in producing conservative simultaneous confidence intervals and hypothesis tests. The current procedure used to compute these upper bounds (Dunn-Sidak method) is based upon making the conservative assumption that the variables are independent.

Three new approaches which give tighter (lower) upper bounds for such probability areas have been developed. The first of these (intersection subtraction) is an improved version of the Bonferonni upper bounds. The second of these methods (conditional multiplicative) requires that the multivariate normal distribution have the MTP-2 property, (~~see~~ Karlin and Rinott for the definition of MTP-2). The third method (conservative independent subunit) is a more complicated form of the conservative assumption of independence among variables.

These three methods are compared theoretically with the following results: (i) The conditional multiplicative, when it can be applied is better than the other two methods. (ii) The intersection subtraction is better than the conservative independent subunit when n is small, but becomes worse than the conservative independent subunit as n becomes larger.

It is shown that these methods can be combined together to give tighter upper bounds than can be given by any one method singly. Different ways of combining the algorithms are examined along with the improvements they give in lowering upper bounds.